

# FIGURE 1A

1 CCCCCTGGAC AGCCGCCCTC TCCTCCAGGC CCGTGGGGT GGCCCTGCAC CGCCGAGCTT CCGGGATGA GGGCCCCCGG TGTTGTCACC CGCGCGGCC

GGGGACCTG TCGCGGGAG AGGAGTCCG GGCACCCGA CCGGACGTG CGGCTCGAA GGGCCCTACT CCGGGGGCC ACACCACTGG GCGCGCGGG

101 CAGGTGCTG AGGACCCCG GCGAGGGCG GAGATGGGG TGCACGGTGA GTACTCGCG GCTGGGGCT CCGCCCCCGG CGGTCCCTG TTTGAGCGGG

GTCCAGCGAC TCCCTGGGC CGTCCGGC CTTACCCC ACGTGCCACT CATGAGCGC CGACCCGGA GGGCGGGCGG GCCCAGGGAC AAACCTCGGCC

1 M G V H

^Exon 1

^Met

201 GATTAGCGC CCGGGCTATT GGCCGGGAGG TGGCTGGGT CAAGGACCGG CGACTTGTC AGGACCCCG AAGGGGAGG GGGTGGGG AGCCTCCACG

CTAAATCGG GCGCCGATAA CCGGCCCTCC ACCGACCCAA GTTCTGGCC GCTGAACAGT TCCTGGGGC TTCCCTCTCC CCCCACCCG TCGGAGGTG

301 TGCCAGCGG GACTTGGGG AGTCTTGGG GATGGGAAA ACCTGACCTG TGAAGGGGAC ACAGTTTGG GGTGAGGGG AAGAAGGTTT GGGGTTCCTG

ACGGTCGCC CTGAACCCC TCAGGAACCC CTACCGTTT TGGACTGGAC ACTTCCCCTG TGTCAAACCC CCAACTCCCC TTCTTCCAAA CCCCCAAGAC

401 CTGTGCCAST GGAGAGGAAG CTGATAAGCT GATAACCTGG GCGTGGAGC CACCACCTAT CTGCCAGAG GNNNTGGTA GCTGGGGTG GGTGTGCAC

GACACGGTSA CCTCTCCTC GACTATTGGA CTATTGGAC CCGGACCTG GTGGTGAATA GACGCTCTCC CNNNNACCAT CGACCCCCAC CCCACACGTG

501 ACGGCAGCAG GATTGAATGA AGGCCAGGGA GGCAGCACCT GAGTGTTCG CAGTGTGGG ACAGGAAGGA CGAGCTGGG CAGAGACGTG GGGATGAAGG

TGCCGTCGTC CTAACCTACT TCCGGTCCCT CCGTGTGGA CTCACGAACG TACCAACCC TGTCCTCTCT GCTCGACCC GTCTCTGCAC CCTACTTCC

601 AAGCTGTCT TCCACAGCCA CCTTCTCCC TCCCGCCTG ACTCTCAGC TGGCTATCTC TTCTAGAATG TCCTGCCCTG CTGTGGCTT TCCTGTCCCT

TTCCACAGGA AGGTGTCGT GGAAGAGGG AGGGCGGAC TGAGAGTCGG ACCGATAGAG AAGATCTTAC AGGACGGACC GACACCGAAG AGGACAGGA

1 E C P A W L W L L S L

^exon 2

701 GGTGTGCTC CTCTGGGCC TCCAGTCTCT GGGGCCCCCA CCAGCCTCA TCTGTACAG CCGAGTCTG GAGAGTACC TCTTGAGGC CAAGGAGCC

CGACAGCGAG GGAGACCCG AGGTACAGGA CCGCGGGGT GGTGGGGT AGACACTGT AGACACTGT GGTCCAGGAC CTCTCCATG AGAACCTCG GTTCTCCCG

13 L S L P L G L P V L G A P P R L I C D S R V L E R Y L L E A K E A

801 GAGATATCA CGGTGAGACC CCTTCCCCAG CACATTCCAC AGAACTCAG CTCAGGGCTT CAGGGAACCTC CTCCCAGATC CAGGAACCTG GCACCTGGTT

CTCTTATAGT GCCACTCTG GGAAGGGGTC GTGTAAGGTG TCTTGAAGTGC GAGTCCCGAA GTCCCTTGAG GAGGTCTAG GTCCTTGAC CGTGAACCAA

46 E N I T

## FIGURE 1B

901 TGGGGTGGAG TTGGGAAGCT AGACACTGCC CCCCTACATA AGAATAAGTC TGGTGGCCCC AAACCATACC TGGAAACTAG GCAAGGAGCA AAGCCAGCAG  
 ACCCACCTC AACCCCTTCCA TCTGTGACGG GGGGATGTAT TCTTATTAG ACCACCGGG TTTGGTATGG ACCTTTGATC CGTTCCCTCGT TTCGGTCTGTC

1001 ATCCTACGGC CTGTGGGCCA GGGCCAGAGC CTTGAGGAC CTTTGACTCC CCGGGCTGTG TGCATTTCAG ACGGGCTGTG CCGAACACTG CAGCTTGAAT  
 TAGGATGCCG GACACCCGGT CCCGGTCTCG GAACTCCCTG GAACTGAGG GGGCCGACAC ACGTAAAGTC TGCCCGACAC GGCTTGTGAC GTCGAACCTTA  
 T G C A E H C S L N  
 ^exon 3

1101 GAGAATATCA CTGTCCCAGA CACCAAGTT AATTCTATG CCTGGAAGAG GATGGAGGTG AGTTCCTTTT TTTTTTTTTT TCCTTTCTTT TGGAGAAATCT  
 CTCTTATAGT GACAGGGTCT GTGGTTTCAA TTAAAGATAC GGACCTTCTC CTACCTCCAC TCAAGGAAAA AAAAAAAGAA ACCTCTTTAGA

1201 CATTTGGCAG CCTGATTTTG GATGAAAGG AGAATGATCG AGGAAAGGT AAAATGGAGC AGCAGAGATG AGGCTGCCCTG GGCGCAGAGG CTCACGTCTA  
 GTAAACGCTC GACTAAAAC CTACTTTCCC TCTTACTAGC TCCCTTTCCA TTTTACCTCG TCGTCTCTAC TCCGACGGAC CCGCGTCTCC GAGTGCAGAT

1301 TAATCCCAGG CTGAGATGGC CGAGATGGGA GAATTGCTTG AGCCCTGGAG TTTTCAGACCA ACCTGGGCAG CATAGTGAGA TCCCCCATCT CTACAAACAT  
 ATTAGGGTCC GACTCTACCG GCTCTACCT CTTAACGAAC TCGGACCTC AAAGTCTGGT TGGACCCGTC GTATCACTCT AGGGGTAGA GATGTTTGTG

1401 TTAATAAAAT TAGTCAGGTG AGGTGGTGCA TGGTGGTAGT CCCAGATATT TGGAAAGCTG AGCGGGGAGG ATCGCTTGAG CCCAGGAATT TGAGGCTGCA  
 AATTTTTTTA ATCAGTCCAC TCCACCACGT ACCACCATCA GGGTCTATAA ACCTTCCGAC TCCGCCCTCC TAGCGAACTC GGTTCCTTAA ACTCCGACGT

1501 GTGAGCTGTG ATCACACCAC TGCACCTCCAG CCTCAGTGAC AGAGTGAGGC CCTGTCTCAA AAAAAGAAAG AATAATGAGG GCTGTATGGA  
 CACTCGACAC TAGTGTGGTG ACGTGAGGTC GGAGTCACTG TCTCACTCCG GGACAGAGTT TTTTCTTTTC TTTTCTTTTC TTATTACTCC CGACATACCT

1601 ATACATTTCAT TATTCATTCA CTCACCTCATT CATTCAATTA TTCAATTCNN NNTTCTTATT GCATACCTCT GTTTGCTCAG CTTGGTGCTT GGGGCTGCTG  
 TATGTAAGTA ATAAGTAAGT GAGTGAGTAA GTAAGTAAGT AAGTAAGNN NNNAGAATAA CGTATGGAGA CAACGAGTC GAACACAGAA CCCCAGCAG

1701 AGGGGCAGGA GGGAGAGGGT GGCATGGGTC AGCTGACTCC CAGAGTCCAC TCCCTGTAGG TCAGGCAGCA GGCCGTAGAA GTCTGGCAGG GCCTGGCCCT  
 TCCCCGTCCT CCCTCTCCCA CCGTACCCAG TCGACTGAGG GTCTCAGGTG AGGACATCC AGTCCGTCGT CCGGCATCTT CAGACCGTCC CGGACCGGGA  
 V R Q Q A V E V W Q G L A L  
 ^exon 4

1801 GCTCTCGGAA GCTGTCTCTG GGGGCCAGGC CCTGTTGGTC AACTCTTCCC AGCCGTGGGA GCCCTGCAG CTGCATGTGG ATAAAGCCGT CAGTGGCCCT  
 CGAGAGCCTT CGACAGGAGC CCCGGTCCG GGACAACCAG TTGAGAGGG TCGCACCTT CCGGACCGTC GACGTACACC TATTTCCGCA GTACCCGGAA

1901 CGGAGCCTCA CCACTCTGCT TCGGGCTCTG GGAGCCGAGG TGAGTAGGAG CGGACACTTC TGCTTGCCCT TTCTGTAAGA AAGGGAGAAG GGTCTTGCTA  
 GCGTCGAGT GGTGAGACGA AGCCCGAGC CCTCGGGTCC ACTCATCTC CCTGTGAAG ACGAACGGGA AAGACATTCT TTCCCTCTTC CCAGAACGAT

# FIGURE 1C

2001 AGGAGTACAG GAACTGTCCG TATTCCTTCC CCTTCTGTGG CACTGCAGCG ACCTCCTGTT TTCTCCTTGG CAGAAGGAAG CCATCTCCCC TCCAGATGCG  
 TCCTCATGTC CTTGACAGGC ATAAGGAAGG GGAAGACACC GTGACGTCCG TGGAGGACAA AAGAGGAACC GTCTTCCTTC GGTAGAGGGG AGGTCTACGC  
 K E A I S P P D A  
 ^exon 5

1

2101 GCCTCAGCTG CTCCACTCCG AACATCACT GCTGACACTT TCCGCAAACT CTTCGGAGTC TACTCCAATT TCCCTCCGGG AAAGCTGAAG CTGTACACAG  
 CGGAGTCGAC GAGGTGAGGC TTGTTAGTGA CGACTGTGAA AGGCGTTTGA GAAGGCTCAG ATGAGGTTAA AGGAGGCCCC TTTCGACTTC GACATGTGTC  
 10 A S A A P L R T I T A D T F R K L F R V Y S N F L R G K L K L Y T G

2201 GGGAGGCCTG CAGGACAGGG GACAGATGAC CAGGTGTGTC CACCTGGGCA TATCCACCAC CTCCCTCACC AACATTGCTT GTGCCACACC CTCCCCCGCC  
 CCCTCCGGAC GTCCTGTCCC CTGTCTACTG GTCCACACAG GTGGACCCCGT ATAGGTGGTG GAGGAGTGG TTGTAACGAA CACGGTGTGG GAGGGGGCGG  
 44 E A C R T G D R O

2301 ACTCCTGAAC CCCGTCGAGG AGCTCTCAG  
 TGAGGACTTG GGGCAGCTCC TCGAGAGTC

## FIGURE 2

1 ATGGGGTGC AGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCTGCT GTCCGCTCCCT CTGGGCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT  
TACCCCCACG TGCTTACAGG ACGGACCGAC ACGGAAGAGG ACAGGAGCGA CAGCGAGGGA GACCCGAGG GTCAGGACCC GCGGGTGGT GCGGAGTAGA  
1 M G V H E C P A W L W L L L S L L S L P L G L P V L G A P P R L I C  
101 GTGACAGCG AGTCCTGGAG AGTACCTCT TGGAGGCCAA GGAGGCCGAG AATATCAGG CCGCTGTGC CGAACACTGC AGCTTGAATG AGAATATCAC  
CACTGTGGC TCAGGACCTC TCCATGGAGA ACCTCCGGTT CCTCCGGCTC TTATAGTCTT GCGGACACG CTTGTGACG TCGACTTAC TCTTATAGTG  
35 D S R V L E R Y L L E A K E A E N I T T G C A E H C S L N E N I T  
201 TGTCOCAGAC ACCAAGTTA ATTTCTATGC CTGGAAGAGG ATGGAGGTCA GGCAGCAGGC CGTAGAAGTC TGGCAGGGCC TGGCCCTGCT CTCGGAAGCT  
ACAGGCTGTG TGGTTTCAAT TAAAGATACG GACCTTCTCC TACCTCCAGT CCGTCGTCCG GCATCTTCAG ACCGTCCCG ACCGGGACGA GAGCCTTCGA  
68 V P D T K V N F Y A W K R M E V R Q Q A V E V W Q G L A L L S E A  
301 GTCCTGCGG GCCAGGCCCT GTTGGTCAAC TCTTCCCAGC CGTGGGAGCC CCTGCAGCTG CATGTGGATA AAGCCGTGAG TGGCCTTCCG AGCCTCACCA  
CAGGACGCC CCGTCCGGGA CAACCAAGTG AGAAGGTCG GCACCTCCG GCACGTGAC GTACACCTAT TTCGGCAGTC ACCGGAAGCG TCGGAGTGGT  
101 V L R G Q A L L V N S S Q P W E P L Q L H V D K A V S G L R S L T T  
401 CTCTGCTTCG GGCTCTGGGA GCCCAGAAGG AAGCCATCTC CCTCCAGAT CCGGCTCCAG CTGCTCCACT CCGAACAAATC ACTGCTGACA CTTTCCGCAA  
GAGACGAGC CCGAGACCCCT CCGGTCTTCC TTCGGTAGAG GGGAGGTCTA CCGCGGAGTC GACGAGGTGA GGCTTGTAG TGACGACTGT GAAAGGCGTT  
135 L L R A L G A Q K E A I S P P D A A S A A P L R T I T A D T F R K  
^Polymorphism in CRL 1609 (A or C). A changes the protein sequence to K.  
501 ACTCTTCCGA GTCTACTCCA ATTTCTCCG GGGAAAGCTG AAGCTGTACA CAGGGGAGGC CTGCAGGACA GGGACAGAT GACCA  
TGAGAAGGCT CAGATGAGGT TAAAGGAGGC CCCTTTCGAC TTCGACATGT GTCCCTCCG GACGTCTGT CCCCTGTCTA CTGGT  
168 L F R V Y S N F L R G K L K L Y T G E A C R T G D R O

### **FIGURE 3**

human MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE

\*\*\*\*\*

chepo MGVHECPAWLWLLLSLLSLPLGLPVLGAPPRLICDSRVLERYLLEAKEAE

human NITTGCAEHCSLNENITVPDTKVNIFYAWKRMEVGQQAVEVWQGLALLSEA

\*\*\*\*\*

chepo NITTGCAEHCSLNENITVPDTKVNIFYAWKRMEVRQQAVEVWQGLALLSEA

human VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRLALGAQKEAISPPD

\*\*\*\*\*

chepo VLRGQALLVNSSQPWEPLQLHVDKAVSGLRSLTTLRLALGAXKEAISPPD

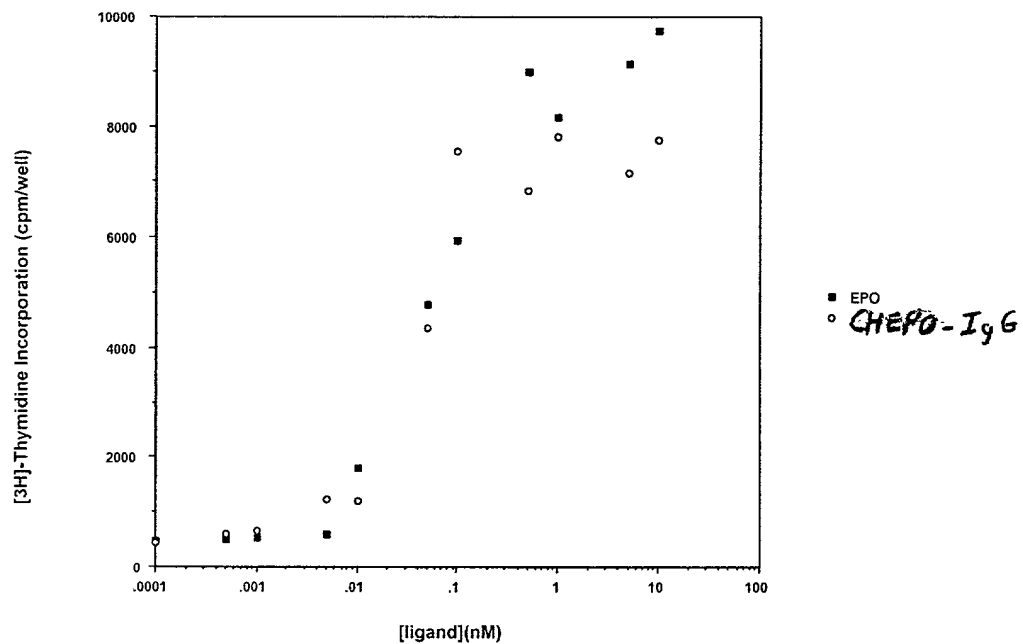
human AASAAPLRTITADTFRKLFRVYSNFLRGKCLKLYTGEACRTGDR

\*\*\*\*\*

chepo AASAAPLRTITADTFRKLFRVYSNFLRGKCLKLYTGEACRTGDR

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 4**



**FIGURE 5**

